DOCUMENT RESUME

ED 121 809 TM 005 237

AUTHOR Ping, Chieh-min; Tucker, Ledyard R.
TITLE Transformation of Both Predictor and Criterion Variables to a Simplified Regression Structure.
PUB DATE [Apr 76]
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS Correlation; Factor Analysis; Matrices; *Multiple Regression Analysis; *Predictor Variables; *Transformations (Mathematics)
IDENTIFIERS Canonical Correlation Analysis; Criterion Variables

ABSTRACT

Prediction for a number of criteria from a set of predictor variables in a system of regression equations is studied with the possibilities of linear transformations applied to both the criterion and predictor variables. Predictive composites representing a battery of predictor variables provide identical estimates of criterion scores as do the original variables. A search for simple structure as in rotation of axes in factor analysis, resulting in transformed regression weights., is demonstrated. This simplified structure has both theoretical and practical importance in application. Illustration of this canonical predictive composite system is demonstrated with three real data sets. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED121809

Transformation of Both Predictor and Criterion Variables To A Simplified Regression Structure

by

Chieh-min Ping
University of Iowa

and

Ledyard R Tucker
University of Illinois at Urbana-Champaign

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

Chieh-min Ping

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Session number D-69

Paper presented at the Annual Meeting of the American Educational Research Association, 1976.

TM005 237

Transformation of Both Predictor and Criterion Variables To a Simplified Regression Structure

Chieh-min Ping
University of Iowa
and
Ledyard R Tucker
University of Illinois at Urbana-Champaign

Prediction for a number of criteria from a set of predictor variables in a system of regression equations is studied with the possibilities of linear transformations applied to both the criterion and predictor variables. Predictive composites representing a battery of predictor variables provide identical estimates of criterion scores as do the original variables. A search for simple structure as in rotation of axes in factor analysis, resulting in transformed regression weights, is demonstrated. This simplified structure has both theoretical and practical importance in application. Illustration of this canonical predictive composite system is demonstrated with three real data sets.

Transformation of Both Predictor and Criterion Variables To a Simplified Regression Structure

Chieh-min Ping[1]
University of Iowa

and

Ledyard R Tucker
University of Illinois at Urbana-Champaign

INTRODUCTION

Canonical correlation analysis is concerned with two functions, one which is taken as the criterion combination. The prediction of the observed attributes as a multidimensional criterion domain and the construction of an estimation function to provide predictions for each observed criterion attribute are recent significant development by Tucker(1973) to be contrasted with the development of criterion combinations by Hotelling. Tucker proposed that the prediction of each criterion attribute from the battery of predictive attributes may be developed by using a separate linear multiple regression equation for each criterion attribute, but that such analyses would not reveal possible basic structure of a prediction system. As Tatsuoka(1973) indicated, "variables simply do not act in the same way when looked at in isolation and when they are considered together with other variables." This disparity is, of course, not confined to predictor variables but holds for criterion variables as well. Tucker pointed out that the space of the predictor attributes could

---

[1]This paper is based on the first author's doctoral dissertation(Ping,1975), written at the University of Illinois at Urbana-Champaign under the supervision of Dr. Maurice M. Tatsuoka. Drs. Ledyard R Tucker, Robert L. Linn, and James L. Wardrop were the other committee members. Grateful acknowledgement is extended to all my committee members.

be partitioned into two uncorrelated subspaces. One is uncorrelated with the criterion attribute domain while the other represents the totality of the relations between the predictor domain and the criterion domain. It is the latter domain that is defined by linear composites of the original predictive attributes. The structure of the predictive subspace and the associate estimation function should provide useful information concerning the prediction situation. A possibility is that transformed composites of both the predictor attributes and the criterion attributes may be found that exhibit simple structure(as in rotation of axes in factor analyses). The regression weights for the composites may be obtained from the regression weights for the original variables. If such a transformation is found, each transformed predictive composite may be interpreted in terms of the criterion attributes for which it is involved in the predictions. As such, the analyses of the regression systems along the lines indicated could reveal simplified structures which would have both theoretical and practical importance in educational research.

The purpose of the present study is (1) to illustrate and investigate the possibility of transformations of the canonical predictive composites which may be considered in terms of factor analytic rotation of axes. If a simple structure can be found in the transformed predictive composites, the transformed composites may be interpreted in terms of the criterion attributes containing high weights on particular transformed composites, and (2) to evaluate and explore use of the canonical predictive component system in the areas of educational application.

THEORETICAL DEVELOPMENT

Consider the situation in which there are two batteries with X containing J predictor attributes and Y containing K criterion attributes. If all measures are taken to be deviations from population means, the following matrix represents the population covariance matrix:

$$\begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix}$$

with C_{xx} for covariances among predictor attributes,

C_{xy} for covariances between predictor and criterion attributes,

C_{yx} for covariances between criterion and predictor attributes, and

C_{yy} for covariances among criterion attributes.

If $\underline{Z}$ is a row vector of measures on M predictive composites obtained by

(1) $$\underline{Z} = \underline{X}\, W,$$

where W is a J x M matrix of weights, then

(2) $$\hat{\underline{Y}} = \underline{Z}\, B',$$

where B' is an M x K matrix of weights.

Let $\underline{\delta}$ be a row vector of differences between $\underline{Y}$ and $\hat{\underline{Y}}$:

(3) $$\underline{\delta} = \underline{Y} - \hat{\underline{Y}}.$$

The weight matrices W and B are to be determined so as to minimize the following loss function:

(4) $$\phi = \xi(\underline{\delta}\, C_{yy}^{-1}\, \underline{\delta}') = \text{Trace}\,(C_{yy}^{-1}\, C_{\delta\delta})$$

where $C_{\delta\delta}$ is a covariance matrix among the differences $\underline{\delta}$.

Possible transformations of a solution in the system are identifiable only within the following class of transformations prior to the minimization of ϕ: Let T by any nonsingular, M x M matrix and

(5) $$W\,T' = \tilde{W},$$

(6) $$B\,T^{-1} = \tilde{B},$$

(7) $$\underset{\sim}{Z}\,T' = \underset{\sim}{\tilde{Z}},$$

then

(8) $$\underset{\sim}{\tilde{Z}} = \underset{\sim}{X}\,\tilde{W}$$

and

(9) $$\underset{\sim}{\hat{Y}} = \underset{\sim}{\tilde{Z}}\,\tilde{B}'.$$

For a convenient derivation, let initial identification of the predictive component system within the foregoing class of transformations, the covariance matrix C_{zz} among predictive composites, be the identity matrix:

(10) $$C_{zz} = I,$$

then

(11) $$C_{\tilde{z}\tilde{z}} = T\,T'.$$

The minimization of ϕ is achieved by setting partial derivatives of ϕ with respect to the weights in B and W equal to zero. Then the solution of B is

(12) $$B = C_{yz}.$$

By the theory of linear composites:

(13) $$C_{\delta\delta} = C_{yy} - C_{yz}\,B' - B\,C'_{yz} - B\,C_{zz}\,B',$$

where

$$C_{zz} = W'\,C_{xx}\,W$$

and

(14) $$C_{yz} = C_{yx} W.$$

Substitution is equation (13) from (10) and (12) yields:

(15) $$C_{\delta\delta} = C_{yy} - C_{yz} C_{yz}'$$

and the loss function in (4) becomes:

(17) $$\phi = \text{Trace } (I - C_{yy}^{-1} C_{yz} C_{yz}')$$
$$= K - \text{Trace } (C_{yz}' C_{yy}^{-1} C_{yz})$$

with the help of equation (14):

(18) $$\phi = K - \text{Trace } (W' C_{yx}' C_{yy}^{-1} C_{yx} W).$$

If W_m is the m'th column vector of W, then introducing λ_m, the corresponding Lagrange multiplier introduced by the condition of equation (10), the solution for W_m will be:

(19) $$(C_{yx}' C_{yy}^{-1} C_{yx} - \lambda_m C_{xx}) W_m = 0$$

This solution is identical with canonical analysis solution. Minimization of ϕ is achieved by selection of the M largest λ_m, the m'th squared canonical correlation and corresponding weight vectors W_m to constitute matrix W. With this solution:

(20) $$\phi = K - \sum_{m=1}^{m} \lambda_m.$$

The loss function should be defined in terms of the criterion space as a whole. It should be not only independent of scale used for measurement of each criterion but also of linear transformations among the criterion attributes. The Mahalanobis distance function used in the above derivations has these properties. The value of this function is designated by ϕ and the function is given by equation (4). Parameters in the predictive system are the dimensionality of the predictive subspace and the entries in the matrices W and B. For any given dimensionality M determined in the predictive subspace, entries in W and B are

to be taken so as to minimize ϕ. The dimensionality M may be determined such that there is no advantage in increasing the number of dimensions and the value of ϕ is not decreased by increasing the number of dimensions.

An extremely important exploration considered next is related to the possibilities of linear transformations of one solution to another solution. To investigate the possibilities of transformations, one proceeds to the point involving equations (5) through (9). The predictive subspace is unique but the coordinate axes for the subspace are not restricted. This is identical to the problem of rotation of axes in factor analysis. Any M x M nonsingular matrix may be regarded as transformation matrix T. Equation (5), (6) and (7) present the transformations to be applied to the matrices $\tilde{W}$ and $\tilde{B}$ and row vector $\tilde{Z}$. With the identification of equation (10), the matrix B turns out to be equal to the covariance matrix of the criterion attributes with the predictive composites as in equation (12), when the predictors being used in a regression system are standardized and uncorrelated. The result given in equation (19) presents conditions for the solutions for columns of the weight matrix W. This solution is identical to the solution for canonical correlation analysis.

Mathematical exploration for the possibilities of transformations to more meaningful composites were given in terms of equations (5) through (11). If transformed composites could be found that exhibit simple weights for prediction of the various criterion attributes, each transformed predictive composite may be interpreted in terms of the criterion attributes for which it is involved in the predictions. A search for simple structure as in rotation of axes in factor analysis

to a desirable position is to approximate a set of rules specified by Thurstone (1947). Actually the simple structure concept is based on the idea of parsimony. It is more of an ideal than a realizable objective. Empirical data have many dependencies that prohibit attaining ideal simple structure but the factors are best located when the produced structure is as simple as possible. After rotation, one hopes that the correlations among the reference axes are never too large, but one can never place any restriction on the correlations among the reference axes without contradicting in some sets of variables the requirements for simple structure.

METHOD

Subjects

The following data were available for a sample of 2,759 beginning freshmen admitted to the University of Illinois at Urbana-Champaign in Fall, 1971. All incoming freshmen are required to take a series of tests given by the Psychological and Counseling Center each fall. These tests consist of two batteries: the Freshman Guidance Exams (FGE) and the Placement and Proficiency Exams (P & P). The FGE includes SCAT and a reading test. The P & P exams include tests of English, math, biology, chemistry and foreign language. Preadmission variables include four ACT individual test scores and four composite high school unit variables which were more stable than individual high school course grades. The four composite high school unit variables are total math, total social science, total science and total language. Only correlations between high school units and pre- and post- admission test scores or between pre- and post-admission

test scores are relevant, since correlations between variables in unrelated areas would not be meaningful.

The second and the third data sets include 3,648 boys and 3,859 girls from the Project TALENT (1967) study. These two groups of subjects were tested when they were in grade 9 in 1960 and retested in grade 12 in 1963 using the same 49 variables. Records were based only on those who took the posttests as well as the pretests.

RESULTS AND DISCUSSION

As indicated in Table 1, the freshman enrollment placement tests, six canonical correlations were significant by the Bartlett approximate chi-square at the .05 level. Since the third squared canonical correlation .1540 was considered as a trivial value, a two dimensional predictive subspace was used in factor analytic rotation. The value of the loss criterion ϕ was 6.7750, equalling eight less the sum of squared canonical correlations for the two dimensions. The initial weight matrix W for obtaining the initial predictive composites is given in Table 2 and the initial weight matrix B and its transformed weights are presented in Table 3.

Possibilities of transformation of the predictive composite system were considered in terms of a two dimensional space for the matrix B. Inspection of Table 3 shows that predictive composite $\tilde{A}$ could be set up uniquely for SCAT Quantitative (2), Advanced Math (7) and Chemistry (8), and predictive composite $\tilde{B}$ could be established for SCAT Verbal (1), Rhetoric (6), Reading Vocabulary (3), Reading Speed (4) and Reading Comprehension (5).

In Table 4, the intercorrelations of transformed composites show that although composites $\tilde{A}$ and $\tilde{B}$ are moderately correlated, they are differentially useful in predictions in that composite $\tilde{A}$ is a unique predictor for quantitativeness whereas composite $\tilde{B}$ is a predictor for verbal achievement.

For the boys' data in Project TALENT, a seven dimensional predictive subspace was used in factor analytic rotation since a value of 0.3932 for the squared canonical correlation was considered trivial. The value of the loss criterion ϕ was 44.0218.

The transformed weights from the predictive composites indicate that the highest loadings on the first dimension include Sports (14),[2] Int. H. S. Math (44), Adv. H. S. Math (45), Arithmetic Comprehension (46), Word Functions (36), and Math (5) in the Information Test. Trivial weights could be established for Spelling (31), Punctuation (33) in the English Test, Social Studies (4) in the Information Test, Abstract Reasoning (42) in General Mental Ability Test, and Arithmetic Reasoning (43) in the Mathematics Test. It seems odd that Sports (14) should load on this dimension which otherwise seems to be pretty much a mathematical factor.

A second transformed predictive composite could be established for which the Practical Knowledge (22) in the Information Test Part II, Mechanics (11) in Part I and Theater (27), Vocabulary II (28) in Part II have high loadings. Disguised Words (30) in the General Mental Ability Test is also loaded high in this dimension.

The third transformed predictive composite could be defined by those variables having high weights including Bible Information (23),

---

[2]Variable number refers to the test variable number in the posttest battery.

Sports (14) and, to a minor degree, Arithmetic Computation (46), Social Studies (4) and Information in the Humanities-Literature (2).

Variables which load high on the fourth dimension include Visualization in three Dimensions (41), Bible Information (23), Abstract Reasoning (42), and Mathematics Information Test (5).

The fifth predictive composite in which significant variables appear include Bible Information Test (23), Word Functions (36), Adv. H. S. Math (45), Mathematics Information Test (15) and Social Studies (4).

The sixth predictive composite could be considered for which Bible Information Test (23), Hunting (24) and Fishing Information Test (25) having significant weights.

The seventh dimension in which significant variables appear includes Bible Information Test (23), Farming (12), Biological Information Test (7), and Humanities-Literature (2).

The resulting transformation matrix and the intercorrelation matrix for the transformed predictive composites show that, among the transformed composites, two correlates highly with three, five and seven; three correlates very highly with five and seven; four with six. Such high intercorrelations among the predictive composites imply that there is very little "individuality" of factors. This is probably the result of the nature of predictor and criterion sets which consist of pre-and post- measures on the same set of variables. This leads to a cycle of high correlations between variables across the two sets. In such a case, it seems that squared canonical correlations will show a gradual decrease all the way through, without any point at which a

sudden drop occurs.

The squared canonical correlations for the girls' data in which the pattern appears is the same as that for the boys'. Hence it is concluded that rotation of regression weights of the predictive composites on girls' data set would not be necessary. In this situation, predictive composites will not give very useful information when the two sets of variables consist of pre- and post- scores on the same set of tests.

CONCLUSION

Canonical correlation analysis is concerned with more than one criterion variable predicted from several independent variables. It is a useful tool in studying and parsimoniously describing the relationship between salient factors in two psychological domains. In contrast to the development of criterion combinations by Hotelling, the canonical predictive composite system is concerned with prediction of individual criterion variables from the battery of predictive variables. Predictive composites representing a battery of predictor variables provide identical estimates of criterion scores as do the original variables. The transformed composites exhibit simple weights for prediction of the various variables and each transformed predictive composite could be interpreted in terms of the criterion variables for which it is involved in the predictions.

Illustrations of the transformations of the canonical predictive composites as indicated in equations (5) through (11) by using factor analytic rotation of axes to a simple structure are demonstrated with three empirical data sets--one with University of Illinois Freshmen

enrollment placement tests on pre- and post- admission variables, two data sets on males and females, respectively, from Project TALENT. Results show that two predictive composites are identifiable for prediction of freshmen placement test at the University of Illinois; the first dimension is a quantitative-scientific one, and the second, a verbal dimension. The boys' data from the TALENT study tentatively yield seven predictive composites, but their interpretations were difficult. It was concluded that for data of this sort (where the two sets are pre- and post-measures on the same test), the method of predictive composites will not lead to fruitful results.

The results of the present study illustrate the usefulness of the canonical predictive component system. Evidence has shown that to establish a smaller number of composite predictor variables than the number of original predictor variables without serious loss in efficiency of the estimates of criterion performances is of prime significance. Therefore, it would be advantageous, in the areas of multivariate analysis, to apply this technique to educational and psychological research.

TABLE 1

Canonical Correlations

Dimension	Chi-square	Df	Squared Canonical Correlation	ϕ	Canonical Correlation
1	6359.1376**	64	.7578	7.2422	.8705*
2	2460.8015**	49	.4672	6.7750	.6835*
3	729.5421**	36	.1540	6.6210	.3924*
4	269.8452**	25	.0618	6.5592	.2486*
5	94.3931**	16	.0230	6.5362	.1517
6	30.3483**	9	.0073	6.5289	.0854
7	10.2366	4	.0037	6.5252	.0608
8	0.0448	1	.0000	6.5252	.0040

*A four dimensional predictive subspace was used in rotation

**Significant at the .05 level

Value of ϕ using 4 dimensions is 6.5592

TABLE 2

Weights for Obtaining Predictive Composites Matrices W and $\tilde{W}$

	Original Weights-Composite			Transformed Weights-Composite		
	Predictor Measure	A	B	Predictor Measure	$\tilde{A}$	$\tilde{B}$
ACT:English	1	0.30552	0.64269	1	0.05738	0.58404
ACT:Math	2	0.46129	-0.70609	2	0.68197	0.05038
ACT:Social Science	3	0.14496	0.29254	3	0.03163	0.27096
ACT:Natural Science	4	0.19634	0.14566	4	0.13182	0.24273
Math	5	0.12773	-0.29679	5	0.22480	-0.03628
Social Science	6	0.04823	0.20088	6	-0.02625	0.14151
Science	7	0.05390	-0.16307	7	0.10830	-0.03407
Language	8	0.01832	0.06545	8	-0.00612	0.04837

TABLE 3

Weights for Predicting Criterion Measures From Predictive Composites Matrices B and $\tilde{B}$

	Original Weights-Composite			Transformed Weights-Composite		
	Criterion Measure	A	B	Criterion Measure	$\tilde{A}$	$\tilde{B}$
SCAT V.	1	0.65451	0.35429	1	0.01700	0.56362
SCAT Q.	2	0.78719	-0.27881	2	0.63251	0.01891
Reading V.	3	0.60384	0.34775	3	-0.00250	0.53952
Reading S.	4	0.34886	0.15690	4	-0.12895	0.44897
Reading C.	5	0.62833	0.23759	5	0.10531	0.44523
Rhetoric	6	0.64763	0.33606	6	0.02937	0.54414
Math	7	0.53945	-0.15453	7	0.40173	0.04711
Chemistry	8	0.39197	-0.11480	8	0.29408	0.03188

TABLE 4

Transformation Matrix T and Intercorrelations of Transformed Composites

Transformation Matrix T

Composite	A	B
$\tilde{A}$	0.93482	-0.35511
$\tilde{B}$	0.86835	0.49595

Intercorrelations of Transformed Composites

Composite	$\tilde{A}$	$\tilde{B}$
$\tilde{A}$	1.00000	0.63564
$\tilde{B}$	0.63564	1.00000

BIBLIOGRAPHY

Darlington, R. B., Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69:161.

Hotelling, H., The most predictable criterion. Journal of Educational Psychology, 1935, 26, 139-142.

Ping, C. M., Application of canonical predictive system to problems in education. Urbana-Champaign, Department of Education, University of Illinois, Unpublished Ph.D. dissertation, 1975.

Shaycoft, M. F., Project TALENT. The high school years: growth in cognitive skills, American Institute for Research and School of Education. Pittsburgh: University of Pittsburgh, 1967.

Tatsuoka, M. M., Multivariate analysis in educational research, In Kerlinger, F. N. (ed.) Review of research in education, Itasca, Illinois: F. E. Peacock, 1973.

Thurstone, L. L., Multiple factor analysis. Chicago: University of Chicago Press, 1947.

Tucker, L. R , A canonical predictive component system. Urbana-Champaign: Department of Psychology, University of Illinois, 1973.